

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/92444/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Petropoulos, Fotios and Kourentzes, Nikolaos 2016. Commentary: two sides of the same coin. Foresight: The International Journal of Applied Forecasting 42 , pp. 37-39. file

Publishers page:

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# **Commentary on “Forecasting: Academia vs. Business”**

## **Two Sides of the Same Coin**

**Fotios Petropoulos and Nikolaos Kourentzes**

Issue 41 of Foresight featured a short commentary by Sujit Singh on the gaps between academia and business. Powered by our focus to produce and disseminate research that is directly applicable to practice, in this commentary we present our views on some of the very useful and interesting points raised by Sujit and conclude with our vision for enhanced communication between the two worlds.

### **ON TRANSLATING ACCURACY TO MONEY**

It is true that the majority of traditional error measures (along with the very widely used in practice MAPE) focus on the performance of point forecasts and their respective accuracy. These are convenient as summary statistics that are context free, but hardly relate to the real decision costs. Therefore, a critical question is how these are translated into business value and how improving forecasting affects utility metrics, such as inventory and backlog costs, customer service level (CSL) and mitigating the bullwhip effect. Fortunately, there is a good bit of research that focuses on such links. Here two very recent examples.

Barrow and Kourentzes (2016) explored the impact of forecast combinations – combining forecasts from different methods -- on safety stocks and found that combinations can lead to reductions compared to using a single ‘best’ forecast. Wang and Petropoulos (2016) evaluated the impact on inventory of base-statistical and judgmentally-revised forecasts. These works show that there is a strong connection between the variance of forecast errors and improved inventory performance.

However, one important point has to be emphasised here: there is limited transparency how forecasts produced by demand planners are translated into ordering decisions by inventory managers. Research typically looks at idealized cases, ignoring the targets and politics that drive inventory decisions. In such cases, the economic benefit of improved forecasts may not reflect organizational realities: forecasting research should pay more attention to the organisational aspects of forecasting.

### **ON WHAT IS GOOD ACCURACY**

Forecast accuracy levels vary across the different industries and horizons. For example a 20% forecast error would be sensible in certain retailing setups, but disastrous in aggregate electricity load forecasting. Short-term forecasting is typically easier, while long-term is more challenging. The nature of the available data is also relevant: fast versus slow moving items; presence of trend and/or seasonality; promotional frequency and so on.

Our approach would be always to benchmark against (i) simple methods, such as naïve or seasonal naïve and (ii) industry-specific (“best practices”) benchmarks. Reporting the improvements in accuracy relative to a these benchmarks helps identify specific problems with the forecasting function and can lead to further refinements. Using relative metrics also overcomes the misplaced focus on what is a good target for percentage accuracy, since these targets do not appreciate the data intricacies that the forecast has to deal with.

## ON AVAILABLE SOFTWARE PACKAGES

Different software packages offer different core features, with some of them specialising in specific families of methods and/or industries. Previously, software vendors were invited to participate in large-scale forecasting exercises (see M3-competition) with the relative rankings of the participating software being available through the original (Makridakis and Hibon, 2000) and subsequent research reports.

In any case, the expected benefits from adopting a software package are a function of data availability, the forecast objective (what needs to be forecast and how long into the future) and the need for automation. Nonetheless, there is need for an up-to-date review and benchmarking of available commercial and non-commercial software packages. Differences exist even in the various implementations of even the simplest methods (such as Simple Exponential Smoothing), with often unknown effects in accuracy. But software packages are important in structuring the forecasting process but vendors often impose their own visions of what is important and these are not often backed up by research. How should one explore the time series at hand? Can we support model selection and specification? How to best incorporate judgemental adjustments?

Our view is that software vendors should provide the tools for users of varying expertise to solve their problems (see comments on customisability by Petropoulos, 2015), but also be explicit about the risks of a solution. Training users is regarded as an important dimension of improving the forecast quality (Fildes and Petropoulos, 2015) as demand planners cannot be replaced by an algorithm. We should not aim for a single solution that will magically do everything and there are always 'horses for courses'.

## ON HIERARCHICAL FORECASTS

Organisations often look at their inventory of data in hierarchies. These can be across products, across markets or across any other classification that is meaningful from a decision making or reporting point of view. Data at different hierarchical levels reveal different attributes of the product history. Although forecasts produced at different hierarchical levels can be translated to forecasts of other levels via aggregation or disaggregation (top-down and bottom-up), the level at which the forecasts are produced will influence the quality of the final forecasts at all the various levels.

Can we know a-priori what is the best level to produce forecasts? Unfortunately, not possible: data have different properties, resulting in different 'ideal levels', but, more importantly, companies have different objectives. Each objective may require different setups.

We believe that the greatest benefit from implementation of hierarchical approaches to forecasting is the resulting reconciliation of forecasts at different decision making levels. The importance of aligning decision-making across levels cannot be understated. More novel techniques allows hierarchies to be forecast and reconciled across different forecast horizons (Petropoulos and Kourentzes, 2014). Recent research (Hyndman and Athanasopoulos, 2014) has demonstrated that approaches that focus on a single levels of the hierarchy, such as top-down or bottom-up, should be replaced with approaches that appropriately combine forecasts (and subsequently information) from all aggregation levels.

It's important to remember that forecasts calculated from data at any level of the hierarchy can be evaluated at all other required levels. One first has to produce the aggregated/disaggregated forecasts and then compare with the actual data points at the respective level.

## FORECASTS ARE USED BY COMPANIES

Research often considers forecasting as an abstract function that is not part of a company or its ecosystem. At the same time, there is ample evidence of the benefits of collaborative forecasting and information-sharing both within the different departments of a company and across the supply chain.

A recent example is provided by Trapero and colleagues (2012) who analyse retail data and show that information sharing between retailer and supplier can significantly improve forecasting accuracy (up to 8 percentage points in terms of MAPE). This research is useful both for modelling in the context of how forecasts are generated and used in organizations.

## A CALL FOR MORE DATA AND CASE STUDIES

Sujit urges production of evidence of “minimum/average/maximum” benefits in different contexts. But current forecasting research has analysed very few data sets. And very few company cases are publicly available. The M1 and M3 competition data sets have been utilised time and again in subsequent studies, so that the results and solutions they derived are susceptible to “over-fitting” and hence not generalisable. Most papers on intermittent demand forecasting make use only of automotive-sales data as well as data sets from the Royal Air Force in the UK. It would be valuable to test our theories and methods on more diverse data sets, but researchers find these are hard to acquire.

We call on practitioners and on vendors to share (after anonymising) empirical data with researchers. The availability of a large number of time series and/or cross-sectional data across a number of industries will increase our understanding of the advantages, disadvantages, and limitations of existing and new forecasting methods, models, frameworks, and approaches.

Researchers are hungry for data while practitioners hunger for solutions to their problems: reducing the barriers will benefit both sides. Still, researchers must appreciate the constraints that limit a company’s willingness to make its data public, and practitioners need to be more proactive in facilitating forecasting researcher.

### References

- Barrow D. and Kourentzes N. (in press) "Distributions of forecasting errors of forecast combinations: implications for inventory management", *International Journal of Production Economics*.
- Fildes R. and Petropoulos F. (2015) "Improving forecast quality in practice", *Foresight: The International Journal of Applied Forecasting* 36, pp. 5–12.
- Hyndman R. and Athanasopoulos G. (2014) "Optimally reconciling forecasts in a hierarchy", *Foresight* 35 (Fall 2014), pp. 42–48.
- Makridakis S. and Hibon M. (2000) "The M3-competition: results, conclusions and implications", *International Journal of Forecasting* 16, pp. 451-476.
- Petropoulos F. & Kourentzes N. (2014) "Improving forecasting via multiple temporal aggregation", *Foresight: The International Journal of Applied Forecasting*, Issue 34 (Summer 2014), pp. 12-17
- Petropoulos F. (2015) "Forecasting Support Systems: ways forward", *Foresight: The International Journal of Applied Forecasting*, Issue 39 (Fall 2015), pp. 5-11.

Wang X. and Petropoulos F. (in press) "To select or to combine? The inventory performance of model and expert forecasts", International Journal of Production Research.

Trapero J.R., Kourentzes N. and Fildes R. (2012) "Impact of Information Exchange on Supplier Forecasting Performance", Omega 40, pp. 738-747.

**Dr Fotios Petropoulos** is an Assistant Professor (Lecturer) at Cardiff Business School of Cardiff University and the Forecasting Support Systems Editor of Foresight. His research expertise lies in behavioural aspects of forecasting and improving the forecasting process, applied in the context of business and supply chain. He is the co-founder, with Nikos, of the Forecasting Society ([www.forsoc.net](http://www.forsoc.net)).

**Dr Nikos Kourentzes** is an Associate Professor (Senior Lecturer) at Lancaster University and a researcher at the Lancaster Centre for Forecasting, UK. His research addresses forecasting issues of temporal aggregation and hierarchies, model selection and combination, intermittent demand, promotional modelling and supply chain collaboration.